FINAL REPORT

MULTISENSOR TRACKING AND RECOGNITION OF ANIMATE AND INANIMATE OBJECTS

J. K. AGGARWAL AND JOYDEEP GHOSH
PRINCIPAL INVESTIGATORS

SEPTEMBER 30, 1994 - DECEMBER 31, 1997

ARO PROPOSAL # 33473-PH CONTRACT # DAAH-04-94-G-0417

THE UNIVERSITY OF TEXAS AT AUSTIN
COMPUTER AND VISION RESEARCH CENTER
AUSTIN, TEXAS 78712

MASTER COPY KEEP FOR REPRODUCTION PURPOSES

REPORT DOCUMENTATION PAGE

Form Approved OMB NO. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 31, 1998	Final Report, January - December 31, 1997

4. TITLE AND SUBTITLE

Multisensor Tracking and Recognition of Animate and Inanimate Objects

5. FUNDING NUMBERS

DAAH-04-94-G-0417

6. AUTHOR(S)

J. K. Aggarwal and Joydeep Ghosh

7. PERFORMING ORGANIZATION NAMES(S) AND ADDRESS(ES)

The University of Texas at Austin
Austin, TX 78712-1084

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park,, NC 27709-2211

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

ARO 33473.21-PH

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12 b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

We report on final results arising from this project, in which we proposed to establish a new paradigm for multisensor tracking and recognition of animate and inanimate objects, fusing a model-based methodology with a neural network-based methodology in an integrated and synergistic manner. Important results are report for the four major project areas: (I) Hybrid ATR systems, (II) Human Motion; (III) Multiple Feature Representation; and (IV) Detection and Tracking of Moving Obstacles in the Path of a Navigating Robot. Major accomplishments include the development of) a hybrid intelligent architecture that exploits the complementary nature of symbolic and connectionist/neural reasoning methodologies for more effective object recognition; (2) a comprehensive mathematical framework to measure the gain in classification performance when several classifiers are combined in a linear fashion; (3) the use of localized gating networks in the mixture-of-experts framework; (4) a Bayesian segmentation framework for textured visual images; (5) a multiple fixed camera system for automatic tracking of human motion in indoor environments; (6) the use of stereo fish-eye lenses for autonomous mobile robot navigation and environment mapping, and (7) an algorithm for moving obstacle detection from a navigating robot.

14. SUBJECT TERMS

multisensor object recognition autonomous navigation
image processing motion analysis automatic target recognition

15. NUMBER IF PAGES

20

16. PRICE CODE

17. SECURITY CLASSIFICATION OR REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

FINAL REPORT

MULTISENSOR TRACKING AND RECOGNITION OF ANIMATE AND INANIMATE OBJECTS

J. K. AGGARWAL AND JOYDEEP GHOSH
PRINCIPAL INVESTIGATORS

SEPTEMBER 30, 1994 - DECEMBER 31, 1997

ARO PROPOSAL # 33473-PH CONTRACT # DAAH-04-94-G-0417

COMPUTER AND VISION RESEARCH CENTER
THE UNIVERSITY OF TEXAS AT AUSTIN
AUSTIN, TEXAS 78712-1084

512/471-3259
HTTP://CVRC.ECE.UTEXAS.EDU

Multisensor Tracking and Recognition of Animate and Inanimate Objects

J. K. Aggarwal & Joydeep Ghosh, Principal Investigators

FINAL REPORT

1. Statement of the Problem Studied.

In this project, we proposed to establish a new paradigm for multisensor tracking and recognition of animate and inanimate objects that fuses a model-based methodology with a neural network-based methodology in an integrated and synergistic manner. Our previous efforts automatic target recognition made apparent the complementary natures of the model-based approach and the neural network-based approaches. Information obtained from one methodology can be used to enhance the image interpretation capabilities of the other. Work under this proposal has centered upon the development of a hybrid approach to ATR that incorporates the positive features of model-based and nonlinear pattern recognition approaches to ATR in a cooperative fashion. Previous research has focused on detecting and recognizing fixed inanimate objects. We have addressed the issues of articulated man-made objects as well as animate objects, to make the hybrid system useful for tracking and recognizing animate objects (humans). Four major project areas were delineated:

(1) Development of a hybrid automatic target recognition system using multisensor fusion that synergistically combines a model-based (symbolic) methodology with a neural network (connectionist) methodology,

(2) Development of a system to recognize and track walking humans from a sequence of images.

(3) Extension of previously developed methods to incorporate multiple feature representations of the object, and

(4) Development of a system to detect and track unexpected moving obstacles that appear in path of a navigating robot.

The focus of this research has been the synergistic implementation of a hybrid ATR system using model-based and neural network-based methodologies for tracking and recognizing fixed man-made objects with articulate motion, as well as animate objects.

Important results are summarized below for the four major project areas:

I. Hybrid ATR Systems.
II. Human Motion
III. Multiple Feature Representation
IV. Detection & Tracking of Moving Obstacles

2. Summary of Important Results of this Project.

Hybrid Systems for Object Recognition.

Prof. Joydeep Ghosh, K. Y. Chang, I. Taha

In an effort to exploit the complimentary nature of symbolic and connectionist/neural reasoning methodologies for more effective object recognition, we conducted an extensive literature survey on this topic and subsequently designed a hybrid intelligent architecture that:

(1) initializes a neural network based on symbolic domain knowledge,

(2) trains this network on known images,

(3) extracts rules from the trained networks, and

(4) combines the resultant expert system with the neural network to provide more reliable decisions with explanation capabilities.

This system was initially tested on a simple problem of controlling the network of dams and reservoirs around Austin. The resulting paper was awarded second prize for the "best application paper" in ANNIE '95. Subsequent work on rule extraction appears in [1]. This is the first end-to-end hybrid system for knowledge based neural networks known to us.

The Hybrid Intelligent Architecture (HIA) exploits the complimentary nature of symbolic and connectionist/neural reasoning methodologies for more effective object recognition. HIA can initialize a neural network based on symbolic domain knowledge, train it on known images, and subsequently extract rules from the trained networks for interpretation. A full journal paper describing HIA and its applications has been published [2]. We have also developed a mechanism for building an expert system (rules plus inference engine) based on the trained network [3]. This mechanism has application in a wide variety of image understanding projects, and provides a fundamental link between data intensive and knowledge intensive approaches.

We have extended the scope of this framework by analyzing non-linear ways of characterizing the data, such as by principal curves, a generalization of principal components. This has the potential of improved and efficient classification, as well as providing a powerful method for describing data, as shown in [4,5].

[1] I. Taha and J. Ghosh, "Three techniques for extracting rules from feedforward networks," in *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol 6, ASME Press, (*Proc ANNIE '96* St. Louis, Nov. 1996), pp. 25-30.

[2] I. Taha and J. Ghosh, "A Hybrid Intelligent Architecture and its Application to Water Reservoir Control", *Int'l Jl. of Smart Engineering Systems*, Vol 1, No. 1, Oct 1997, pp. 59-75.

[3] I. Taha and J. Ghosh, "Evaluation and Ordering of Rules Extracted from Feedforward Networks", *Proc. Int'l Conference on Neural Networks*, Houston, TX, June 1997, pp. 408-13.

[4] K.-Y. Chang and J. Ghosh, "Principal Curve Classifier - A Nonlinear Approach to Pattern Classification", *1998 IEEE Intl. Conf. on Neural Networks*, Anchorage, May 1998.

[5] K.-Y. Chang and J. Ghosh, "Principal curves for nonlinear feature extraction and classification," *Proc. SPIE Conf. on Applications of Artificial Neural Networks in Image Processing III*, SPIE Proc. Vol. , San Jose, CA, Jan 1998.

Evaluation of Object Recognition Systems.

J. K. Aggarwal, A. Mitiche, D. Nair

Although ATD/R is a well-studied subject and a wide range of algorithms, paradigms and systems have been proposed to solve this problem, there is an unfulfilled need for performance criteria to aid the comparison of various systems (algorithms), provide the means to predict their performance in given scenarios, and understand the reliability/robustness of a system and its components.

We have proposed a formal and systematic methodology for the evaluation and comparison of recognition systems [1, 2] based on statistical and algorithmic indicators. Statistical indicators measure the significance of performance difference and provide a ranking of performances when the difference is significant. We use the Kruskal-Wallis "H" test for this purpose. Algorithmic indicators include both (1) ordinary space and time complexity, which measures computation cost, and (2) various performance curves in variables of error and test data sample size, which expand on the definition of performance and underline the special status of the test-data sample, which may not be size adequate. The asymptotic behavior of these curves compensates for the small size of data samples when performance is measured.

To demonstrate the usefulness of this methodology, we compared the performance of a number of AOR systems which differed only in their method of pattern category assignment. These methods are well documented and commonly used in pattern classification, particularly in automatic target recognition (ATR): the multilayer perceptron, Kohonen's self-organizing memory, Carpenter and Grossberg's ART-2 network and the nearest neighbors classifier [3]. The systems were compared on their performances on a database of about 1000 images of 6 different objects.

[1] D. Nair, A. Mitiche and J. K. Aggarwal, "On Comparing the Performance of Object Recognition Systems," *Proc. IEEE International Conference on Image Processing*, pp. II-631-II-634.

[2] Dinesh Nair, Amar Mitiche and J. K. Aggarwal, "A Methodology for Ranking Performances," Computer and Vision Research Center Technical Report TR-97-06-111.

[3] A. Mitiche and J. K. Aggarwal, "Pattern Category Assignment by Neural Networks and Nearest Neighbors Rule," *Intl. J. of Pattern Recognition and Artificial Intelligence*, Vol. 10, No., 5, 1996, pp. 393-408.

Multilearner Systems: Classifier Ensembles.

Prof. Joydeep Ghosh, K. Bollacker, K. Tumer

We have developed a comprehensive mathematical framework that quantifies the gains achieved when several classifiers are combined in a linear fashion [1]. This analysis inspects the reduction in the variance of the decision boundaries around the optimum (Bayes) boundaries, as more classifiers are added. A nice byproduct of this research is a new way of estimating the Bayes error rate, a fundamental quantity. This technique gives much more accurate results than current methods, and is computationally inexpensive [2]. Different ways of training individual classifiers so that the gains of combining are enhanced have also been studied, with extensive simulations [3]. In addition, a novel algorithm for fast classification of foveated images was developed [4], and a novel feature selection approach based on mutual information has been developed [5].

We have further enhanced our mathematical framework that quantifies the gains achieved when several classifiers are combined. In particular, the effect of order statistics has been analyzed, leading to significantly more robust ensembles [6]. Further experimentation has also shown the power of estimating the Bayes error rate, a fundamental quantity based on observing the performance of an ensemble and comparing with that of individual classifiers.

A new study has been to incorporate the knowledge embodied in existing classifiers that may be only weakly related to the new task, to improve, both computationally and performance-wise, the current task of interest [7]. This provides a basis for understanding how existing ATR algorithms and classifiers can be used to help building the next generation systems. The nice scalability properties (with respect to the number of existing classifiers used) of the integration mechanism have been demonstrated [8,9].

[1] I. Taha and J. Ghosh, "A Hybrid Intelligent Architecture and its Application to Water Reservoir Control", Int'l Jl. of Smart Engineering Systems, Vol 1, No. 1, Oct 1997, pp. 59-75.

[2] K. Tumer and J. Ghosh, "Bayes Error Rate Estimation Through Classifier Combining," *Proc. Intl. Conf. on Pattern Recognition,* Vienna, Austria, Aug. 1996, pp. IV:695-99.

[3] V. Ramamurti and J. Ghosh, "Structurally Adaptive Modular Networks for Non-stationary environments", *Proc. Intl. Conf. on Pattern Recognition,* Vienna, Austria, Aug. 1996, pp. II:704-708.

[4] T. Kuyel and J. Ghosh, "Multiresolution Nearest Neighbor (MNN) Classifier: A fast algorithm for the classification of images", Proc. IASTED Conf. on Signal and Image Processing *(SIP-96),* Orlando, Nov. 11-14, 1996, pp 122-25.

[5] K. Bollacker and J. Ghosh, "Linear Feature Extractors based on Mutual Information," *Proc. Intl. Conf. on Pattern Recognition,* Vienna, Austria, Aug. 1996, pp. IV:720-24.

[6] K. Tumer and J. Ghosh, "Classifier Combining through Trimmed Means and Order Statistics", 1998 IEEE Intl. Conf. on Neural Networks, Anchorage, May 1998.

[7] K. Bollacker and J. Ghosh, "A Scalable Method for Classifier Knowledge Reuse", Proc. Int'l Conference on Neural Networks, Houston, TX, June 1997, pp. 1474-79.

[8] K. Bollacker and J. Ghosh, "Knowledge reuse in multiclassifier systems", Pattern Recognition Letters, 18 (11-13), Nov 1997, 1385-90.

[9] K. Bollacker and J. Ghosh, "On The Design of Supra-Classifiers for Knowledge Reuse", 1998 IEEE Intl. Conf. on Neural Networks, Anchorage, May 1998.

Flexible Modular Networks for Nonstationary Environments.

Prof. J. Ghosh, W. S. Chaer, A. Nag, V. Ramamurti

Because many ATR problems are highly context-dependent, it is desirable to design systems that can work in different environments and even in drastically changing environments. We are examining modular neural networks that can respond to both slowly changing and abruptly changing environments. For slow changes, on-line algorithms are used, while abrupt changes are countered by switching among different modules. In our initial work in this area, we developed a fast learning algorithm for modular networks [1,2] with promising experimental results and defined a framework for using a bank of adaptive Kalman filters [3]. The results were extended to online situations relevant to modeling non-stationary environments as described in [4,5]. Model selection was tackled by developing new algorithms that add/delete experts as training takes place. Our work on using localized gating networks in the mixture-of-experts framework was completed, and will appear as an invited talk [6]. We have also shown that regularization can be used to improve performance, provide confidence measures and performance metrics [7,8].

We also have investigated the resource allocation network, again adaptively changing the network structure in response to computational demands, and suited for non-stationary environments. Initial results appear in [9].

In addition, we have used the relationship between the width of localized basis functions and scale-space image processing to yield innovative and powerful image filtering/restoration techniques [10]. This work received the Best Conference Paper Award at ANNIE '97, and postulates a dynamic interpretation of edges.

[1] V. Ramamurti and J. Ghosh, "Advances in using Hierarchical Mixture of Experts for Signal Classification," *Proc. ICASSP* -96, Atlanta, GA, May 1996, pp. 3569-72.

[2] I. Taha and J. Ghosh, "Three techniques for extracting rules from feedforward networks", in *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol 6, ASME Press, (*Proc ANNIE '96* St. Louis, Nov. 1996), pp. 25-30

[3] W. S. Chaer, R. H. Bishop and J. Ghosh, "A Mixture-of-Experts Framework for Adaptive Kalman Filtering," *IEEE Trans. Systems, Man and Cybernetics*, Vol. 27B, No. 3, June 1997. pp. 452-64.

[4] V. Ramamurti and J. Ghosh, "Flexible Modular Architecture for Changing Environments" in *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol 6, ASME Press, (*Proc ANNIE '96* St. Louis, Nov. 1996), pp. 25-30.

[5] V. Ramamurti and J. Ghosh, "Structural Adaptation in Mixture of Experts," *Proc. Intl. Conf. on Pattern Recognition,* Vienna, Austria, Aug. 1996, pp. II:704-708.

[6] V. Ramamurti and J. Ghosh, "Localized Gating in Mixture of Experts Network", (invited paper) *Proc. SPIE Conf. on Applications and Science of Computational Intelligence*, SPIE Proc. Vol. , Orlando, April 1998.

[7] V. Ramamurti and J. Ghosh, "Improved Generalization in Localized Mixture of Experts Networks" in *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol 7, ASME Press, (Proc ANNIE '97), Nov 1997, pp. 5-10.

[8] V. Ramamurti and J. Ghosh, "Regularization and Error Bars for the Mixture of Experts Networks" *Proc. Int'l Conference on Neural Networks*, Houston, TX, June 1997, 221-226.

[9] A. Nag and J. Ghosh, "Flexible resource allocating network for noisy data", *Proc. SPIE Conf. on Applications and Science of Computational Intelligence*, SPIE Proc. Vol. , Orlando, April 1998.

[10] J. Eledath, J. Ghosh and S.V. Chakravarthy, "Image Enhancement using Scale-based Clustering Properties of the Radial Basis Function Network" in *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol 7, ASME Press, (Proc ANNIE '97), Nov 1997, pp. 447-452.

II. TRACKING HUMAN MOTION

Prof. J. K. Aggarwal, Dr. Amar Mitiche, Qin Cai

Although the problem of recognizing and tracking rigid objects has been well studied, the recognition and tracking of moving, non-rigid objects such as the human body has received much less attention [1]. Our objective was to develop a multisensor sensor system for detecting the presence of humans in a secured environment by integrating visual and thermal information using neural network techniques. The use of two sensing modalities would enable the system to obtain fast and accurate detection with few false alarms. In pursuit of this goal, our work initially evolved from studying human walking in a fixed camera [2,3] to tracking non-background objects in a moving camera [4]. In [5], we used a moving camera with a substantial degree of rotational freedom to follow the subject of interest automatically. However, this strategy was still limited in the amount of area it could cover and was too complicated for real time applications, as it requires estimating the motion of both the viewing system and the subject of interest. A comprehensive framework for a multiple, fixed camera system has been developed to capture sequences of synchronized monocular grayscale images. Multivariate Gaussian models are applied to find the most likely matches of human subjects between consecutive frames taken by cameras mounted in various locations. The system consists of three major components, (1) Single View Tracking, (2) Multiple View Tracking, and (3) Automatic Camera Switching. Bayesian classification schemes based on motion analysis of human features are used to track (spatially and temporally) a subject image of interest between consecutive frames. The automatic camera switching module predicts the position of the subject along a spatial-temporal domain and then selects the camera which provides the best view and which requires the least switching to continue tracking. Limited degrees of occlusion are tolerated within the system. Tracking is based upon the images of upper human bodies captured from various viewing angles, and non-human moving objects are excluded using Principal Component Analysis (PCA). Experimental results from real data show the robustness of the algorithm and its potential for real time applications [6].

[1] J.K. Aggarwal, Q. Cai, W. Liao and B. Sabata, "Non-rigid Motion Analysis: Articulated and Elastic Motion," *Computer Vision and Image Understanding*, in press.

[2] J. A. Webb and J. K. Aggarwal. "Structure from motion of rigid and jointed objects," *Artificial Intelligence*, Vol. 19, pp 107-130, 1982.

[3] A. G. Bharatkumar, K. E. Daigle, M. G. Pandy, Q. Cai, and J. K. Aggarwal. "Lower limb kinematics of human walking with the medial axis transformation," *Proc. of IEEE Computer Society Workshop on Motion of Non-Rigid and Articulated Objects*, Austin, Texas, November 1994, pp 70-76.

[4] Q. Cai, A. Mitiche, and J. K. Aggarwal. ìTracking human motion in an indoor environment," *Proc. 2nd Intl. Conf. on Image Processing*, Washington, D.C., October 23-26, 1995, pp 215-218.

[5] Q. Cai and J. K. Aggarwal, "Tracking Human Motion Using Multiple Cameras," Proc. 13th International Conference on Pattern Recognition, Vienna, Austria, August 25-30, 1996, pp. C: 68-72.

[6] Q. Cai and J. K. Aggarwal, "Automatic Tracking of Human Motion in Indoor Scenes Across Multiple, Synchronized Video Streams," *Proc. 1998 International Conference on Computer Vision*, Bombay, India, January 4-7, 1998.

III. MULTIPLE FEATURE REPRESENTATION.

Robust Automatic Target Recognition In Second Generation Forward Looking Infra-Red (FLIR) Images.

Prof. J. K. Aggarwal, Dinesh Nair

Automatic Target Detection and Recognition (ATD/R) is one of the key components of present and future defense weapon systems to be used in autonomous missions. The ATD/R process entails identifying the location of the target(s) in a scene and recognizing its identity and pose. The process takes as input a collection of data (typically in the form of an image) from a sensor or multiple sensors, preprocesses the data to remove noise effects and enhance target information, detects all possible target locations in the image, and, finally, recognizes the identity and pose of the target. ATD/R, therefore, involves processing at all levels of machine vision: lower-level vision, as with edge detection and image segmentation; mid-level vision, as with representation and description of pattern shape, and feature extraction; and higher-level vision, as with pattern category assignment.

The ATD/R process, which deals with problems such as recognizing a target (a battle tank, for example) from one or more images, is a challenging application for the general techniques developed in image processing, image understanding and computer vision. The problem is challenging because:

(1) The targets appear in complex environments.

(2) The targets may appear along with other less important objects, or there may be occlusions and cluttering.

(3) The target signatures vary depending on the surrounding background and environmental conditions, and are generally not repeatable.

One of the basic limitations in ATD/R efforts is associated with imaging deficiency. Early Forward Looking Infrared (FLIR) based sensors did not determine the absolute thermal temperature of the targets. This introduced an image variability that compounded the problems mentioned earlier. Even with the evolution of FLIR technology over the last decade, there are still limitations in the form of high false alarm rates due to background clutter and occlusions from terrain and vegetation. This has motivated the examination of other sensors such as millimeter wave (MMW) radar, synthetic aperture radar (SAR), laser radar (LADAR) and visible electro-optical (EO) sensors and their integration. However, in this research we limit ourselves to images obtained from FLIR sensors.

Given the challenging nature of the ATD/R problem due to the complexity of the imaged scenes, the use of *a priori* information is critical to solving the problem. The fundamental needs of an ATD/R system using a single sensing modality include:

(1) The use of *a priori* knowledge to assist the ATD/R process at all stages.

(2) A good representation of targets and backgrounds, which warrants the use of signatures that are descriptive and robust to target and environmental variations.

(3) The use of a compact set of maximally discriminating features to represent the target. This helps to keep the size of the system as well as the computational complexity manageable.

(4) The ability to adapt dynamically to the changing environment.

This research presents robust algorithms for the ATD/R process. Specifically, the problem addressed in this research is as follows:

> "...develop algorithms for automatic detection and recognition of targets in second generation FLIR images. The algorithms [will] emphasize the use of scene and sensor information, concentrate on few highly discriminatory representations of the target and perform robustly in cluttered environments."

We now review our research in this area.

A Focused Target Segmentation Paradigm.

A new set of algorithms was developed and tested for the segmentation of targets from second generation FLIR images [1]. An initial detection algorithm is used to identify regions in the image that are candidate locations of objects by accurately modeling the background using Weibull functions. A focused analysis of each candidate target location is then performed to get an accurate representation of the target boundary. A region-growing procedure, driven by the underlying probability distribution of the background and modulated by local shape changes of the target, is used to get an initial estimate of the target shape, which is then refined using the salient edge information in the image to arrive at a more accurate representation of the target boundary. A computationally efficient and flexible method to incorporate the salient edge information into the region boundary has been developed by formulating it as a Bayes classification problem. Finally, each detected area is classified as a man-made or natural object by feature models developed using competitive modular neural networks. Geometric and FLIR intensity-based features extracted from the target areas are used for the classification. This segmentation paradigm has been successfully used on images from the Huli9306_sig subset of the Comanche dataset.

[1] D. Nair and J. K. Aggarwal, "A Focused Target Segmentation Paradigm," 1996 European Conference on Computer Vision, April 14-18, 1996, Cambridge, England, pp. I-579-588.

Hierarchical, Modular Architectures for Object Recognition by Parts.

In this part of the research, we have studied the problem of object recognition from the perspective of recognition by object parts. Our methodology is based on a hierarchical, modular structure (HMS) for object recognition, in which the type of recognition performed differs from level to level. Each level is made up of modules, where each module is an *expert* on a particular part of an object, that is, each module is specifically trained to recognize one part of an object. In general, the lowest level of the hierarchy identifies the class of the vehicle (e.g., tanks vs. trucks), while higher levels use information from the lower levels, as well as features extracted from the original object parts, for classification within each class (e.g., an M-60 tank vs. an M-1 tank). The object features used at each level of the hierarchy are determined by their usefulness for the objective at that level. That is, at the lowest level, we use only features that are salient (i.e., the features are unique to the current class of vehicles and can be readily used to discriminate it from other classes of vehicles). Identification of the salient parts could be done either by humans or automatically.

Each modular *expert* is trained to recognize the part under different viewing angles and transformations (translation, scaling and rotation). When presented with an input object part, each *expert* provides a measure of confidence of that part belonging to the object that the *expert* represents. These confidence estimates are used at the higher levels for more refined classification. Bi-directional interactions between modules at the same recognition levels and at different recognition levels are present. By allowing top-down expectations, faster learning and better recognition performances can be achieved.

The modular *experts* may be built using different techniques. For example, the HMS structure can be constructed using a Bayesian approach, where each module represents the conditional probability density function of a part, and the outputs of these modules are then used to estimate the posterior probability of the input part belonging to a specific object. Another approach to building the HMS is to use modular neural networks for each *expert*, where each neural network expert is trained to recognize a specific object part. The outputs of these networks are then combined hierarchically to obtain the final object recognition.

We have developed a Bayesian methodology for recognition of 2D objects using their parts, based on a hierarchical, modular structure for object recognition [1]. Input to the recognition system are the detected/segmented objects in second generation Forward Looking Infra-Red (FLIR) images. The segmented objects are obtained using a detection algorithm that models the background using Weibull functions to identify candidate target locations in the image [2], [3]. A two-stage focused analysis of each candidate target location is then performed to get an accurate representation of the target boundary. A region-growing procedure is used to get an initial estimate of the target region, which is then combined with salient edge information in the image to arrive at a more accurate representation of the target boundary. The region and edge integration is done using a novel method that uses a Bayes' minimum risk classification approach. Finally, to reduce the false alarm rate, a higher level interpretation module is used to classify the detected areas as man-made or natural objects using geometric and FLIR-intensity based features extracted from the target. A 100% detection rate with a false alarm rate of 5% was obtained when the segmentation method was tested on 200 images from the HULI9306_SIG subset of the COMANCHE data set.

For the object recognition system, the lowest level consists of classifiers that are trained to recognize the class of the input target, while at the next level, classifiers are trained to recognize specific targets. At each level, the targets are recognized by their parts, and thus each classifier is made up of modules, each of which is an *expert* on a specific part of the target. Each modular *expert* is trained to recognize one part under different viewing angles and transformations [4]. In Bayesian theory, this information can be determined by finding the target and pose that maximize the a posteriori probability knowing the likelihood and prior models. This Bayesian realization of the methodology has already been developed, in which the expert modules represent the probability density functions of each part, modeled as a mixture of densities to incorporate different views (aspects) of each part. The presence of a specific target in the image is decided by accumulating evidence from the part *experts* for that target. The inputs to the HMS are features extracted from the different parts of a target that have to be recognized. These inputs are presented sequentially to the system, that is, each part is seen by the system one after the other.

For the experimental results presented here, two distinct sets of data were used, one for training the system and the other for testing the system. The **training set** consisted of six targets from three classes (tanks, trucks and armored personnel carriers). For each target, a total of 72 views was considered (0-360_ at 5_ intervals). Since in these images, the targets were imaged from roughly the same elevation, the 72 views captured the variations of the targets completely. The **testing set** consisted of images of the same vehicles used for training but obtained under different viewing conditions and varying segmentation outputs. For each target there were 72 images, giving a total of 432 images to test the system. We divided the testing set into three categories based on the segmentation results, namely, "good," "faulty," and "occlusion."

(1) **Good.** This category consisted of 230 images where the segmentation results looked like the training set.

(2) **Faulty.** This category consisted of 170 images, where the images were segmented poorly (i.e., in many cases, the segmentation results included parts of the background, especially tracks left by the vehicle), and

(3) **Occlusion.** This category consisted of 32 images in which the targets were occluded.

Table I gives the recognition results obtained from these experiments. In these experiments, it was enforced that a decision always be made (i.e., a target was recognized even though the resulting maximum probability of recognizing the target was very low). The first column gives the type of segmentation, the second column denotes how many of the targets were classified correctly into their general class (lowest level of recognition). The third column gives the recognition rate of the targets. The overall recognition rate was 90.05% (391/432). Recognition errors in the "Good" segmentation category occur when the viewing aspect is 0_ or close to it (i.e., a front or back view of the target), and different targets tend to look alike. Some errors in the "Faulty" segmentation category arose due to certain background objects that looked like parts of a target.

Table I

Target Segmentation	Class Recognition	Target Recognition	Misclassified
Good	223 / 230 = 96.9%	216 / 230 = 93.9%	14 / 230 =6.1%
Faulty	155 / 170 = 91.18%	151 / 170 = 88.8%	19 / 170 = 11.2%
Occlusion	23 / 62 = 87.05%	24 / 32 = 75.0%	8 / 32 = 25.0%

Given the hierarchical modular structure, each module in the framework can be improved and expanded upon individually. Methodologies other than Bayesian, such as Neural Networks and Expert Systems will be explored in the context of object recognition. An initial comparison of these methodologies was presented in a related study [5]. The developed system has been tested on six targets so far and we will test the system on many more targets and also under different environmental conditions. A substantial portion of the Comanche Dataset will be used for this purpose. We will also focus on improving the performance to the system by studying further the statistical characteristics of the targets and the background in an imaged scene.

[1] D. Nair and J.K. Aggarwal, "Hierarchical, modular architectures for object recognition by parts," *Proc. International Conference on Pattern Recognition,* Vienna, Austria, August 25-30, 1996, pp. 601-606.

[2] D. Nair and J.K. Aggarwal, "A focused target segmentation paradigm," *Proc. European Conference on Computer Vision*, April 1996, pp. I:579--588.

[3] D. Nair and J. K. Aggarwal, "Region/Edge-based Target Segmentation of FLIR Images Modeled by Weibull/Gaussian Distributions," Computer and Vision Research Center Technical Report TR-97-05-110.

[4] D. Nair and J.K. Aggarwal, "Robust automatic target recognition in second generation FLIR Images," *Proc. IEEE Workshop on Applications of Computer Vision,* Sarasota, FL, December 1996, pp. 194-201.

[5] J.K. Aggarwal, J. Ghosh, D. Nair, and I. Taha, A comparative study of three paradigms for object recognition: Bayesian statistics, neural networks, and expert systems," in *Advances in Image Understanding (a Festschrift for Azriel Rosenfeld)*, K. Bowyer and N. Ahuja, Eds., IEEE Computer Society Press, 1996, pp. 241-262.

Hypergraph Representation for Finding the Surface Correspondence and Estimating Motion from a Pair of Range Images.

J. K. Aggarwal, Bikash Sabata

A novel procedure for finding the surface correspondence and estimating the motion transformation of a moving object from a sequence of images was developed using a hypergraph representation [1]. The two scenes are modeled as hypergraphs and the hyperedges are matched using a sub-graph isomorphism algorithm. The hierarchical representation of hypergraphs not only reduces the search space significantly, but also facilitates the encoding of topological and geometrical information used to direct the search procedure. Results obtained from pairs of range images show that the algorithm is robust and performs well in the presence of occlusions and incorrect segmentations. Motion transformation between image frames is computed using the planar and quadric surface pairings. A least squares minimization procedure is formulated that estimates the best motion transform, subject to the constraints of rigid motion. Motion computation for linear feature pairs thus becomes tractable because the rotation and translation computations become independent of one another. However, this is not true for quadric surfaces. The equation to be minimized is highly non-linear and the uniqueness of solution cannot be guaranteed. The solution obtained computes the motion by extracting unique linear features from the quadric surfaces and using them to compute the motion transformation. The main contribution of the work is a surface-based framework for motion estimation from a sequence of range images. The primary issues of correspondence and motion computation are formulated and solved in terms of surface descriptions.

[1] "Surface Correspondence and Motion Computation from a Pair of Range Images," Bikash Sabata and J. K. Aggarwal, *Computer Vision and Image Understanding*, Vol. 63, No. 2, March 1996.

Bayesian Segmentation Framework for Textured Visual Images

J. K. Aggarwal, S. Shah

Segmentation is an integral part of the computer vision and image analysis paradigm, in which regions of interest are identified and extracted for subsequent processing. In most image analysis applications, the first step is to partition the image into regions that satisfy certain constraints. The segmentation process uses these constraints to construct homogeneous regions and smooth boundaries. Region homogeneity can be determined by using properties such as gray level intensity, color, texture, etc. In images of natural terrain, texture provides significant information that can be used to characterize local image behavior. Texture segmentation involves the identification of the uniform textured regions in an image. Many techniques have been proposed for texture analysis, from simple statistical models to estimate probability density functions, to adaptive filters, and intensity and texture measures, etc. Similarities in the extracted features define homogeneity for a region. Once a potential region is localized, it is extracted from the background. The success of higher level recognition subsystems depends upon the accuracy of the segmentation results. For this reason, segmentation is the most widely studied area of computer vision.

We pose the segmentation problem as a classification of pixels in homogeneous regions using a Bayesian framework [1]. The problem may also be posed as that of texture classification, where Gabor Wavelets are used to extract relevant features. The features obtained are coarse clustered to obtain the approximate region labelings. Each cluster is considered to be suboptimal, with missing data, and thus the parameters are estimated using the Expectation-Maximization (EM) algorithm. Final segmentation is then performed recursively while maximizing the posterior probability for each region.

Tested on real scene images, our segmentation algorithm showed clear distinction between the four regions of a test composite image, with errors seen only at the region boundaries due to the windowing effect during feature extraction using Gabor wavelets. In applying the algorithm to an image containing tactical targets, we extended our initial segmentation by performing region refinement, in which a region-growing procedure is used to analyze the classified texture regions by incorporating measures of local shape characteristics to obtain smooth boundaries and region homogeneity. Results are presented on visual images from the MIT VisTex Texture Database. Manmade object segmentation is also illustrated by extending the basic framework to incorporate characteristics of manmade objects.

[1] Shah, Shishir and J. K. Aggarwal, "A Bayesian Segmentation Framework for Texture Visual Images," *Proc. 1997 Computer Vision and Pattern Recognition Conference*, June 17-19, 1997, Puerto Rico, pp. 1014-1020.

IV. ROBOT NAVIGATION

Autonomous Mobile Robot Navigation and Environment Representation Using Stereo Fish-Eye Lens Camera.

J. K. Aggarwal, Dinesh Nair

A wide variety of approaches and algorithms have been developed in recent years for the autonomous navigation of mobile robots. An autonomous mobile robot navigation system that uses stereo fish-eye lenses has been developed for navigation in an indoor structured environment. The system estimates the three-dimensional (3D) position of significant features in the scene, and by estimating its relative position, navigates through narrow passages and makes turns at corridor ends. Fish-eye lenses provide a large field of view, which helps in imaging objects close to the robot and in making smooth transitions in the direction of motion [1,2]. Calibration is performed for the lens-camera setup and the distortion is corrected to obtain accurate quantitative measurements. A vision-based algorithm that uses the vanishing points of extracted segments from a scene in a few 3D orientations provides an accurate estimate of the robot orientation. This is used, in addition to 3D recovery via stereo correspondence [3], to maintain the robot motion in a purely translational path as well as to remove the effects of any drifts from this path from each acquired image. Horizontal segments are used as a qualitative estimate of change in the motion direction and vertical segment correspondence provides for precise 3D information about objects close to the robot. Assuming detected linear edges in the scene as the boundaries of planar surfaces, the 3D model of the scene is generated. The system is implemented for RoboTex, the mobile robot at our center and tested in a structured environment. Finally, the construction of computer aided design (CAD) models of a structured scene as imaged by the stereo fish-eye lenses is investigated to determine the merits of such a system over some of the other implementations for navigation and modeling of the scene discussed [4]. It is seen that an environment such as a corridor is mainly composed of linear edges with particular orientations in 3D. The linear edges are boundaries of opaque planar patches, such as the floor, ceiling, walls, etc. Repeated estimation and updates of the depth map via correspondence of the linear edges at each step allows the robot to make decisions regarding the navigable path. As the estimated depth has lower uncertainty close to the robot [3], it is possible to navigate in narrow environments. Knowing the 3D segment representations of the robot's environment, the CAD model can be generated by considering isolated segments under guidelines discussed in [5,6].

[1] Shishir Shah and J. K. Aggarwal, "A Simple Calibration Procedure for Fish-Eye Lens Camera, *Proc. 1994 IEEE Intl. Conf. on Robotics and Automation*, San Diego, California, May 1994. pp. 3422-3427.

[2] Shishir Shah and J. K. Aggarwal, "Intrinsic Parameter Calibration Procedure for a (High-Distortion) Fish-Eye Lens Camera with Distortion Model and Accuracy Estimation," *Pattern Recognition*, Vol. 26, No. 1, February 1996.

[3] Shishir Shah and J. K. Aggarwal, "Depth Estimation Using Stereo Fish-Eye Lenses." *Proc. Intl. Conf. on Image Processing*, Austin, Texas, November 1994, pp. 740-744.

[4] Shishir Shah and J. K. Aggarwal, "Modeling Structured Environments Using Robot Vision, in *Recent Developments in Computer Vision*, Lecture Series in Computer Science: Recent Advances in Computer Vision, pp. 113-128, 1995.

[5] Shishir Shah and J. K. Aggarwal, "Autonomous Mobile Robot Navigation Using Fish-Eye Lenses," *Proc. of the 3rd Intl. Computer Science Conference*, December 13-15, 1995, Hong Kong.

[6] Shishir Shah and J. K. Aggarwal, "Mobile Robot Navigation and Scene Modeling Using Stereo Fish Eye Lens System" *Machine Vision and Applications*, 1997, accepted for publication (CVRC TR-97-08-113.)

Robot Self-Location Using Visual Reasoning Relative to A Single Target Object.

J. K. Aggarwal and M. Magee

We have developed a computationally straightforward method to determine the location of a camera that is mounted on a robot [1]. The procedure consists of observing a sphere upon which two great circles have been circumscribed and computing the location and orientation of the camera based upon projections of the primary features in the image plane. Distance from the standard mark is based on the size of the projected sphere, whereas location of the camera is based on the displacements of the great circles relative to the center of the projected sphere. The procedure for determining the location of the camera involves solving only linear equations and is computationally simple. Experimental results confirmed the utility of the method with both simulated data and actual images.

[1] "Robot Self Location Using Visual Reasoning Relative to a Single Target Object," M. Magee and J. K. Aggarwal, *Pattern Recognition*, Vol. 28, No. 2, pp. 125-134, 1995.

Mobile Robot Self-Location Using Model-Image Feature Correspondence.

J. K. Aggarwal and R. Talluri

We have developed an approach to solving the problem of establishing reliable and accurate correspondence between a stored 3-D model and its 2D image in the context of autonomous mobile robot navigation in an outdoor urban, man-made environment [1]. The environment is assumed to consist of polyhedral buildings, and 3D descriptions of the lines constituting the buildings' rooftops are assumed to be given as a world model. The robot's position and pose are estimated by establishing correspondence between the straight line features extracted from the images acquired by the robot and the model features. The correspondence problem is formulated as a two-stage constrained search problem. Geometric visibility constraints are used to reduce the search space of possible model-image feature correspondences. Techniques for effectively deriving and capturing these visibility constraints from the given world model were developed. The position estimation technique was robust and accurate even in the presence of occlusions, errors in the feature detection and incomplete model descriptions.

[1] "Mobile Robot Self-Location Using Model-Image Feature Correspondence," R. Talluri and J. K. Aggarwal, *IEEE Trans. on Robotics and Automation*, Vol. 12, No. 1, February 1996, pp. 63-77.

Moving Obstacle Detection from a Navigating Robot.

Prof. J. K. Aggarwal, Dinesh Nair

We have developed a system that detects unexpected moving obstacles that appear in the path of a navigating robot and that estimates the relative motion of the object with respect to the robot [1]. The system is designed for a robot navigating in a structured environment with a single wide-angle camera. The system uses polar mapping to simplify the segmentation of the moving object from the background. The polar mapping is performed with the Focus of Expansion (FOE) as the center. A vision-based algorithm that uses the vanishing points of segments extracted from a scene in a few 3-D orientations provides an accurate estimate of the robot orientation. This is used to maintain the motion of the robot along a purely translational path and also used to subtract the effects of any drifts from this path from each image acquired by the robot. By doing so, the determination of the FOE is simplified. In the transformed space, a qualitative estimate of moving obstacles is obtained by detecting the vertical motion of edges extracted in a few specified directions. Relative motion information about the obstacle is then obtained by computing the time_to_impact between the obstacles and the robot from the radial component of the optical flow.

The system was implemented and tested on an indigenously fabricated autonomous mobile robot, RoboTex [2]. RoboTex is a 1.5 meter tall, tetherless mobile robot, weighing about 150 kg. The robot's subsystems are comprised of a TRC Labmate base and rigid metal frame; a fast onboard UNIX workstation to digitize video images and control the robot; a camera and digitizer; an I/O system; and a power supply to enable completely autonomous operation. The Labmate base can carry 90 kg of equipment at speeds of up to 1 meter/second, and accelerations of 10 cm s^{-2}. We use it at 40 cm/s and 5 cm s^{-2} to avoid wheel slippage and remove motion blur. The right and left driving wheels are mounted on a suspension for good floor contact. Passive casters in each corner ensure stability. The Labmate controller processes measurements from the right and left odometer to update the 2D position and heading of the robot. We found that, provided that accelerations were reasonable, the odometric readings were reliable. The rigidity of the frame is important since the transformation between the coordinate systems of the camera and the robot must be calibrated precisely.

The effectiveness of the qualitative estimate of the motion was tested extensively on numerous runs in typical building corridors. The obstacles encountered in these tests were moving humans and opening doors. The following steps were accomplished to detect obstacles and estimate the time_to_impact:

1. Acquire an image from the camera sensor.
2. Acquire orientation of robot.
3. Derotate image.
4. Perform polar mapping of the image.
5. Extract horizontal and angular edges from the image.
6. Detect qualitatative motion.
7. Compute optical flow.
8. Determine time_to_impact.

Using a PA-RISC based HP-735 workstation running at 99 MHz, when step 2 is obtained directly from the robot odometry, the system was able to detect moving obstacles at 100 ms/frame. If the orientation of the robot is acquired from the vanishing point of significant lines in the image, an additional 7 ms was required. The system thus is proved useful as a cueing mechanism for moving obstacles in structured, indoor environments (mainly corridors). To our knowledge, there are no documented experimental demonstrations of systems that detect moving obstacles from moving platforms with comparable performance to this system.

[1] "Moving Obstacle Detection from a Navigating Robot," D. Nair and J. K. Aggarwal, *IEEE Trans. on Robotics and Automation,* in press.

[2] X. Lebegue and J. K. Aggarwal, "Robotex: An autonomous mobile robot for precise surveying," *Proc. Intl. Conf. On Intelligent Autonomous Systems,* Pittsburgh, PA, pp. 460-469, February 1993.

3. Publications and Technical Reports.

Aggarwal, J. K. and Shishir Shah, "Object Recognition and Performance Bounds," In *Lecture Notes in Computer Science: Image Analysis and Processing*, pp. 343-360, Edited by Alberto Del Bimbo, Springer Verlag, 1997. (Proceedings of the 9th Intl. Conference on Image Analysis and Processing, Florence, Italy, September 1997.)

Aggarwal, J.K., Q. Cai, W. Liao and B. Sabata, "Non-rigid Motion Analysis: Articulated and Elastic Motion," *Computer Vision and Image Understanding*, in press.

Bollacker K. and J. Ghosh, "Mutual Information Feature Extractors for Neural Classifiers," *Proc. IEEE Intl. Conf. On Neural Networks*, Washington, DC, June 1996, pp. 1528-33.

Bollacker, K. and J. Ghosh, "Linear Feature Extractors based on Mutual Information," *Proc. Intl. Conf. on Pattern Recognition,* Vienna, Austria, Aug. 1996, pp. IV:720-24.

Bollacker, K. and J. Ghosh, "A Scalable Method for Classifier Knowledge Reuse," *Proc. Int'l Conference on Neural Networks*, Houston, TX, June 1997, pp. 1474-79.

Bollacker, K. and J. Ghosh, "Knowledge reuse in multiclassifier systems," *Pattern Recognition Letters*, 18 (11-13), Nov 1997, 1385-90.

Bollacker, K. and J. Ghosh, "On The Design of Supra-Classifiers for Knowledge Reuse," *IEEE Intl. Conf. on Neural Networks*, Anchorage, May 1998, accepted for publication.

Cai, Qin, A. Mitiche and J. K. Aggarwal, "Tracking Human Motion in an Indoor Environment," *Proc. IEEE International Conference on Image Processing*, Washington, D.C., October 23-26, 1995, pp. I215-I218.

Cai, Q. and J. K. Aggarwal, "Tracking Human Motion Using Multiple Cameras," Proc. 13th International Conference on Pattern Recognition, Vienna, Austria, August 25-30, 1996, pp. C: 68-72. .

Cai, Q. and J. K. Aggarwal, "Automatic Tracking of Human Motion in Indoor Scenes Across Multiple Synchronized Video Streams," *Proc. 1998 International Conference on Computer Vision*, Bombay, India, January 4-7, 1998.

Chaer, W. S., R. H. Bishop and J. Ghosh, "A Mixture-of-Experts Framework for Adaptive Kalman Filtering," *IEEE Trans. Systems, Man and Cybernetics*, Vol. 27B, No. 3, June 1997. pp. 452-64. .

Chang, K.-Y. and J. Ghosh, "Principal Curve Classifier - A Nonlinear Approach to Pattern Classification," *IEEE Intl. Conf. on Neural Networks*, Anchorage, May 1998, accepted for publication.

Chang, K.-Y. and J. Ghosh, "Principal curves for nonlinear feature extraction and classification," *Proc. SPIE Conf. on Applications of Artificial Neural Networks in Image Processing III*, SPIE Proc. Vol. , San Jose, CA, Jan 1998..

Chang, Yuhlin and J. K. Aggarwal, "Line Correspondences from Cooperating Spatial and Temporal Grouping Processes for a Sequence of Images," *Computer Vision and Image Understanding*, Vol. 67, No. 2, August 1997, pp. 194-201.

Chang, Yuh-lin and J. K. Aggarwal, "Representing and Estimating 3-D Lines," *Pattern Recognition*, Vol. 28, No. 8, pp. 1181-1190, August 1995.

Dhond, Umesh and J. K. Aggarwal, "Stereo matching in the presence of narrow occluding objects using Dynamic Disparity Search, " *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 7, pp. 719-724, July 1995.

Eledath, J., J. Ghosh and S.V. Chakravarthy, "Image Enhancement using Scale-based Clustering Properties of the Radial Basis Function Network," in *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol 7, ASME Press, (Proc ANNIE '97), Nov 1997, pp. 447-452.

Kuyel, T. and J. Ghosh, "A fast space localized computation of the outputs of a Gabor Filter bank", *Proc. IASTED Conf. on Signal and Image Processing* (SIP-95.), Las Vegas, Nov. 20-23, 1995.

Kuyel, T. and J. Ghosh, "Multiresolution Nearest Neighbor (MNN) Classifier: A fast algorithm for the classification of images," *Proc. IASTED Conf. on Signal and Image Processing (SIP-96),* Orlando, Nov. 11-14, 1996, pp 122-25.

Magee, Michael, and J. K. Aggarwal, "Robot Self Location Using Visual Reasoning Relative to a Single Target Object," *Pattern Recognition*, Vol. 28, No. 2, pp. 125-134, February 1995.

Mitiche, Amar and J. K. Aggarwal, "Pattern Category Assignment by Neural Networks and Nearest Neighbors Rule," *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 10, No. 5, 1996, pp. 393-408.

Nag, A. and J. Ghosh, "Flexible resource allocating network for noisy data," *Proc. SPIE Conf. on Applications and Science of Computational Intelligence*, SPIE Proc. Vol. , Orlando, April 1998, accepted for presentation.

Nair, D. and J. K. Aggarwal, "Hierarchical, Modular Architectures for Object Recognition by Parts," *Proc. 13th International Conference on Pattern Recognition*, Vienna, Austria, August 25-30, 1996, pp. A: 601-606.

Nair, D. and J. K. Aggarwal, "A Focused Target Segmentation Paradigm," *Proc. 1996 European Conference on Computer Vision,* April 14-18, 1996, Cambridge, England.

Nair, D., A. Mitiche and J. K. Aggarwal, "On Comparing the Performance of Object Recognition Systems," *Proc. IEEE International Conference on Image Processing*, Washington, D.C., October 23-26, 1995, pp. II-631-II-634.

Nair, Dinesh and J. K. Aggarwal, "A Focused Target Segmentation Paradigm," *Proc. of the Fourth European Conference on Computer Vision*, Cambridge, England, April 14-18, 1996, pp. I: 579-588.

Nair, Dinesh and J. K. Aggarwal, "Moving Obstacle Detection from a Navigating Robot," *IEEE Transactions on Robotics and Automation*, accepted for publication

Nair, Dinesh and J. K. Aggarwal, "Region/Edge-based Target Segmentation of FLIR Images Modeled by Weibull/Gaussian Distributions," Computer and Vision Research Center Technical Report TR-97-05-110.

Nair, Dinesh and J. K. Aggarwal, "Robust Automatic Target Recognition in Second Generation FLIR Images," *Proc. IEEE Workshop on Applications of Computer Vision*, December 2-4, 1996, pp. 194-201.

Nair, Dinesh, Amar Mitiche and J. K. Aggarwal, "A Methodology for Ranking Performances," Computer and Vision Research Center Technical Report TR-97-06-111.

Ramamurti, V. and J. Ghosh, "Advances in using Hierarchical Mixture of Experts for Signal Classification," *Proc. ICASSP*-96, Atlanta, GA, May 1996, pp. 3569-72.

Ramamurti, V. and J. Ghosh, "Flexible Modular Architecture for Changing Environments" in *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol 6, ASME Press, (*Proc ANNIE '96* St. Louis, Nov. 1996), pp. 25-30.

Ramamurti, V. and J. Ghosh, "Localized Gating in Mixture of Experts Network," (invited paper) *Proc. SPIE Conf. on Applications and Science of Computational Intelligence*, SPIE Proc. Vol. , Orlando, April 1998, accepted for presentation.

Ramamurti, V. and J. Ghosh, "Structural Adaptation in Mixture of Experts," *Proc. Intl. Conf. on Pattern Recognition,* Vienna, Austria, Aug. 1996, pp. II:704-708.

Ramamurti, V. and J. Ghosh, “Improved Generalization in Localized Mixture of Experts Networks,” in *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol 7, ASME Press, (Proc ANNIE '97), Nov 1997, pp. 5-10.

Ramamurti, V. and J. Ghosh, “Regularization and Error Bars for the Mixture of Experts Networks,” *Proc. Int'l Conference on Neural Networks*, Houston, TX, June 1997, pp. 221-226.

Sabata, Bikash and J. K. Aggarwal, "Surface Correspondence and Motion Computation from a Pair of Range Images," *Computer Vision and Image Understanding*, Vol. 63, No. 2, March 1996, pp. 232-250.

Shah, Shishir and J. K. Aggarwal, "A Bayesian Segmentation Framework for Texture Visual Images," *Proc. 1997 Computer Vision and Pattern Recognition Conference*, June 17-19, 1997, Puerto Rico, pp. 1014-1020.

Shah, Shishir and J. K. Aggarwal, "Autonomous Mobile Robot Navigation Using Fish Eye Lenses," presented at the 3rd International Computer Science Conference, December 13-15, 1995, Hong Kong. Conference proceedings published as *Image Analysis Applications and Computer Graphics*, Lecture Notes in Computer Science, Volume 1024, Roland T. Chin et al, editors, pp. 9-16, 1995, Springer Verlag.

Shah, Shishir and J. K. Aggarwal, "Intrinsic Parameter Calibration for a (High Distortion) Fish Eye Lens Camera with Distortion Model and Accuracy Estimation," *Pattern Recognition*, Vol. 26, No. 11, 1996, pp. 1775-1788.

Shah, Shishir and J. K. Aggarwal, "Modeling Structured Environments Using Robot Vision," in *Recent Developments in Computer Vision*, S. Z. Li, D. P. Mital, E. K. Teoh and H. Wang, Eds., Springer-Verlag, 1996, pp. 113-128.

Shah, Shishir and J. K. Aggarwal, "Modeling Structured Environments Using Robot Vision," *Proceedings of the 2nd Asian Conference on Computer Vision*, December 5-8, 1995, Singapore, pp. I297-I304.

Shah, Shishir and J. K. Aggarwal, "Mobile Robot Navigation and Scene Modeling Using Stereo Fish-eye Lens System, " *Machine Vision and Applications*, Vol. 10, 1997, pp. 159-173.

Shah, Shishir, J. K. Aggarwal, Jayakkrishnan Eledath and J. Ghosh, "Multisensor Integration for Scene Classification: An Experiment in Human Form Detection, " *Proc. IEEE Signal Processing Society International Conference on Image Processing*, Santa Barbara, CA, October 1997, vol. 2, pp. 199-202.

Taha, I. and J. Ghosh, "A Hybrid Intelligent Architecture and its Application to Water Reservoir Control", *Int'l Jl. of Smart Engineering Systems*, Vol 1, No. 1, Oct 1997, pp. 59-75.

Taha, I. and J. Ghosh, "Controlling Water Reservoirs Using a Hybrid Intelligent Architecture", in *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol. 5, ASME Press, (Proceedings of the 1995 Conference on Artificial Neural Networks in Engineering, St. Louis, MO, Nov. 1995), pp. 63-68.

Taha, I. and J. Ghosh, "Three techniques for extracting rules from feedforward networks", in *Intelligent Engineering Systems Through Artificial Neural Networks*, Vol 6, ASME Press, (*Proc ANNIE '96* St. Louis, Nov. 1996), pp. 23-28.

Taha, I. and J. Ghosh, "Evaluation and Ordering of Rules Extracted from Feedforward Networks," *Proc. Int'l Conference on Neural Networks*, Houston, TX, June 1997, pp. 408-13.

Talluri, Raj and J. K. Aggarwal, "Mobile Robot Self-Location Using Model-Image Feature Correspondence," *IEEE Transactions on Robotics and Automation*, Vol. 12, No. 1, February 1996, pp. 63-77.

Tumer K. and J. Ghosh, "Error Correlation and Error Reduction in Ensemble Classifiers" *Connection Science*, Special issue on Combining, Vol. 8, No. 3/4, Dec 1996, pp. 385-404.

Tumer K. and J. Ghosh, "Estimating the Bayes Error Rate through Classifier Combining," *Proc. Intl. Conf. on Pattern Recognition,* Vienna, Austria, Aug. 1996, pp. IV:695-99.

Tumer, K. and J. Ghosh, "Theoretical Foundations of Linear and Order Statistics Combiners for Neural Pattern Classifiers," *IEEE Trans. on Neural Networks*, in press.

4. Participating Scientific Personnel and Advanced Degrees arising from this Work.

Participating Scientific Personnel

J. K. Aggarwal	Faculty Summer Research
Joydeep Ghosh	Faculty Summer Research
Amar Mitiche	Senior Research Fellow
Louis Hornung	Research Associate
Kurt Bollacker	Graduate Research Assistant
Qin Cai	Graduate Research Assistant
Biao Lu	Graduate Research Assistant
Dinesh Nair	Graduate Research Assistant
Bryan W. Stiles	Graduate Research Assistant
Ismail Taha	Graduate Research Assistant
Viswanath Ramamurti	Graduate Research Assistant
Shishir Shah	Graduate Research Assistant
Xiaowei Wang	Graduate Research Assistant
Stephen Cheung	Undergraduate Research Assistant
Phillip K. S. Chu	Undergraduate Research Assistant
Arvind Siota	Undergraduate Research Assistant
Faraz Syed	Undergraduate Research Assistant
Nancy Truong	Undergraduate Research Assistant

Advanced Degrees.

Dinesh Nair, "Robust Automatic Target Recognition in Second Generation Forward Looking Infrared Images, " present address: National Instruments, Austin, Texas (Graduated December 1996).

Shishir Shah, "Autonomous Mobile Robot Navigation in Structured Environments Using Stereo Fish-Eye Lens Camera, " M.S. Thesis, August 1996. Currently pursuing Ph.D.

Ismail Taha, "A Hybrid Intelligent Architecture for Revising Domain Knowledge", May 1997. (Trilogy Systems, Austin).

Brian Stiles, "Nonlinear Memory Dynamic Neural Networks for Spatio-Temporal Classification", May 1997 (JPL, Pasadena).

Viswanath Ramamurti, "Structurally Adaptation and Generalization in Modular Neural Networks", August, 1997 (Technology Resources Inc (research division of SBC Corp.), Austin)

5. List of Inventions.

None.